

Hyperparameter tuning, Batch Normalization, Programming Frameworks

Quiz, 10 questions

8/10 points (80%)

Congratulations! You passed!	Next Item	
1/1 points		
1. If searching among a large number of hyperparameters, you should try values in random values, so that you can carry out the search more systematically and not False?		
True		
False		
Correct		
X 0/1 points		
2. Every hyperparameter, if set poorly, can have a huge negative impact on training, hyperparameters are about equally important to tune well. True or False?	and so all	
O True		
This should not be selected No. We've seen in lecture that some hyperparameters, such as the learning rat than others.	e, are more critical	
C False		
1/1		
points		
3. During hyperparameter search, whether you try to babysit one model ("Panda" st models in parallel ("Caviar") is largely determined by:	rategy) or train a lot of	

Whether you use batch or mini-batch optimization

Hyperpara The presence of local minima (and saddle points) in your neural network Hyperparameter tuning, Batch Normalization, Programming Frameworks amount of computational power you can access

8/10 points (80%)

Quiz, 10

questions Corre	
	The number of hyperparameters you have to tune
×	0 / 1 points
	nink $oldsymbol{eta}$ (hyperparameter for momentum) is between on 0.9 and 0.99, which of the following is the nended way to sample a value for beta?
	1 r = np.random.rand() 2 beta = r*0.09 + 0.9
	1 r = np.random.rand() 2 beta = 1-10**(- r - 1)
	1 r = np.random.rand() 2 beta = 1-10**(- r + 1)
0	1 r = np.random.rand() 2 beta = r*0.9 + 0.09
This s	hould not be selected
~	1/1 points
5. Finding	good hyperparameter values is very time-consuming. So typically you should do it once at the

start of the project, and try to find very good hyperparameters so that you don't ever have to revisit tuning them again. True or false?

\bigcirc	True
0	False

Hyperparameter tuning, Batch Normalization, Programming Frameworks

8/10 points (80%)

Quiz, 10 questions



1/1 points

6.

In batch normalization as presented in the videos, if you apply it on the lth layer of your neural network, what are you normalizing?



 $z^{[l]}$

Correct

-		г
()	$W^{\scriptscriptstyle 1}$

$$b^{[l]}$$

$$\bigcirc a^{[l]}$$



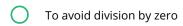
1/1 points

7.

In the normalization formula $z_{norm}^{(i)}=\frac{z^{(i)}-\mu}{\sqrt{\sigma^2+\varepsilon}}$, why do we use epsilon?

()	To have a more accurate normalization
١		/	TO Have a Hibre accurate Horrialization

To speed up convergence



Correct

In case μ is too small



1/1 points

8.

Which of the following statements about γ and β in Batch Norm are true?

They can be learned using Adam, Gradient descent with momentum, or RMSprop, not just with gradient descent.

Hyperparameter tuning, Batch Normalization, Programming Frameworks

8/10 points (80%)

Quiz, 10 questions

questior	is		
	There is one global value of $\gamma\in\Re$ and one global value of $\beta\in\Re$ for each layer, and applies to all the hidden units in that layer.		
lln-s	elected is correct		
011-3	elected is correct		
Corr	They set the mean and variance of the linear variable $z^{m{l}} l$ of a given layer.		
	The optimal values are $\gamma=\sqrt{\sigma^2+\varepsilon}$, and $\beta=\mu$.		
Un-s	Un-selected is correct		
	eta and γ are hyperparameters of the algorithm, which we tune via random sampling.		
Un-s	elected is correct		
05			
~	1 / 1 points		
9.			
	raining a neural network with Batch Norm, at test time, to evaluate the neural network on a new		
	le you should:		
ожар			
0	Perform the needed normalizations, use μ and σ^2 estimated using an exponentially weighted average across mini-batches seen during training.		
Correct			
	Skip the step where you normalize using μ and σ^2 since a single test example cannot be normalized.		
	If you implemented Batch Norm on mini-batches of (say) 256 examples, then to evaluate on one test example, duplicate that example 256 times so that you're working with a mini-batch		

Use the most recent mini-batch's value of μ and σ^2 to perform the needed normalizations.



the same size as during training.

Correct

Frambayou	rameter tuning, Batch Normalization, Programming rks ngf these statements about deep learning programming frameworks are true? (Check all that	8/10 points 80%) It apply)
	Deep learning programming frameworks require cloud-based machines to run.	
Un-s	selected is correct	
	Even if a project is currently open source, good governance of the project helps ensure that it remains open even in the long term, rather than become closed or modified to benefit cone company.	
Corre	rect	
	A programming framework allows you to code up deep learning algorithms with typically flines of code than a lower-level language such as Python.	fewer



